420-TP-003-001

SDPS End to End Scenarios Overview

Technical Paper

Technical Paper—Not intended for
formal review or government approval.

March 1995

Prepared Under Contract NAS5-60000

RESPONSIBLE ENGINEER

Ron Williamson /s/	4/4/95
Ron Williamson, HRL Research Project Manager EOSDIS Core System Project	Date

SUBMITTED BY

Stephen Fox /s/	6/2/95
Steve Fox, SDPS Manager EOSDIS Core System Project	Date

Hughes Applied Information Systems
Landover, Maryland

This page intentionally left blank.

Abstract

This technical paper provides an overview of the ECS End to End scenarios presented at the February, 1995 Science Data Processing System (SDPS) Preliminary Design Review (PDR). The scenarios address the functionality and data archives available at both the Release A and B time frames. The scenarios provide both a user view and a system view of the SDPS and related Communication and System Management Segment (CSMS) functionality concentrating primarily on application services to be developed by the SDPS development team. A separate user view is characterized in each of the three representative scenarios as the Earth Scientist's view, the Instrument Scientist's view and the System Administrator and Operator's view. The separate Scientist views capture the roles that an individual scientist may play at any point in the life-cycle of EOSDIS related research and does not imply that scientist are responsibly for exclusively one role or the other.

Keywords: SDPS, Scenarios, System, Release A, Release B, Science View, Operations View, System View

This page intentionally left blank.

Contents

Abstract

1. Introduction

2. End to End Scenarios Overview

3. Instrument Scientist Scenario

4. Earth Scientist Scenario

5. Administrative / Operations Scenario

6. Suggested Reading

Figures

Acronyms

This page intentionally left blank.

1. Introduction

1.1 Purpose

The purpose of this technical paper is to provide a brief overview of the end to end system scenarios presented at the SDPS PDR during February, 1995. This paper characterizes each of three scenarios presenting distinct views of the ECS system: Earth Scientist View, Instrument Scientist View, and Administration/Operations View. Each of the scenarios described presents both the user's view and the system view. The user's view is described in terms of a realistic science or system administration context and the system view is described in terms of the system architecture components defined in the SDPS SDS and the SDPS PDR material.

1.2 Organization

This paper is organized as follows:

Section 1.0 Introduction and executive summary of the End to End scenarios.

Section 2.0 End to End Scenario Overview

Section 3.0 Earth Scientist Scenario

Section 4.0 Instrument Scientist Scenario

Section 5.0 System Administration / Operations Scenario

Section 6.0 Suggested Reading

1.3 Review and Approval

This Technical Paper is an informal document approved at the Office Manager level. It does not require formal Government review or approval; however, it is submitted with the intent that review and comments will be forthcoming.

Questions regarding technical information contained within this Paper should be addressed to the following ECS :

- ECS Contacts
  - Ron Williamson, HRL Project Manager, 301.925.0340, william@eos.hitc.com
  - Steve Fox, SDPS Manager, 301.925.0346, sfox@eos.hitc.com

Questions concerning distribution or control of this document should be addressed to:

Data Management Office
The ECS Project Office
Hughes Applied Information Systems,
1616A McCormick Dr.,
Landover, MD 20785

2. End to End Scenarios Overview

2.1 Scenario Introduction

This document provides an overview of the end to end scenarios for Release A and Release B functionality of the EOSDIS Core System (ECS) design. The overview is intended as an introduction to the ECS architecture and design at the PDR level and is not intended as a detailed discussion of the design or operations concept. Separate, voluminous documents describe the ECS design at the PDR level in great detail and are available via the Internet and World Wide Web (WWW) server at the following Uniform Resource Locator (URL) http://edhs1.gsfc.nasa.gov/.

The end to end scenarios were developed in a collaborative manner using the resources of the ECS Science Data Processing Segment (Ron Williamson), the ECS Science Office (Tom Dopplick), Land Processes science expertise (Bill Emery and Dan Baldwin, University of Colorado), Earth's Radiation Budget science expertise (Bruce Barkstrom's team at LaRC), and the Maintenance and Operations expertise in the ECS M&O organization (Stan Dunn) and the SDPS Data Server Subsystem team (Mark Huber and Mike Burnett).

This document is intended as an executive summary bridge between the overview of the system architecture presented at the ECS SDR and PDR introduction and those aspects of the system described in detail in the ECS design documentation. The documentation includes the System Design Specification and the detailed PDR specification presented at the ECS PDR and included in the SDPS Segment/Element Design Specification document.

This technical document also is intended as a science oriented view of the ECS Architecture. The key subsystems addressed include the Client, Interoperability, Data Management, Data Server, Data Processing, Planning, and Ingest. Each of these subsystems is described in detail in both the PDR presentation material and the SDPS Segment/Element Design Specification document. Each subsystem is described via an object model, a set of detailed scenarios, event traces, and an interface specification. Each of the end to end scenarios uses a subset of the object model to defined key subsystem components used in the scenario, a subset of the event traces of the detailed scenarios, and a subset of the key interfaces to each subsystem.

2.2 ECS Scenario Context

The ECS as a system must support a diverse set of users (disciplines, expertise, objectives, methods, tools) who are geographically distributed and use widely varying computational and networking capabilities. An inherent part of the scientific research method is that increased understanding leads to changes in scientific algorithms. The proposed architecture must support the fact that the algorithms and interfaces will change many times over the life of the system.

From a data perspective the system contains a diverse set of data types and representations, an extremely large data volume (TB/day and PB archives), complex data product interdependencies, and a requirement for long-term archival of data as a “national resource."

The ECS architecture must also support system attributes such as scalability, evolvability, reliability, maintainability, availability in order to be useful to Earth Science research. The system must support evolving technologies , changing data sources, and varying investigative approaches. The ECS must also fit within, and indeed provide reusable components to, a federation of systems and sites to meet broader national and international Earth Science objectives (GCDIS, UserDIS contexts).

Figure 2-1 depicts the ECS via the core functionality of the system. From left to right the system must support a wide range of data and algorithm sources via the Ingest subsystem. Once ingested the system must support the production and archiving of valuable data products for subsequent retrieval by Earth Science researchers and eventually a wider user community. The system must also support users in finding and access relevant data collections and services of interest to them via information discovery and management services. The system must also support the publishing of research results (whether data, algorithms, or research papers) and the subsequent local analysis of those results at user sites. The system also has a goal of supporting and indeed facilitating collaborative multi-disciplinary research in support of the Mission to Planet Earth objective of an integrated inter-disciplinary model of the Earth's Environment.

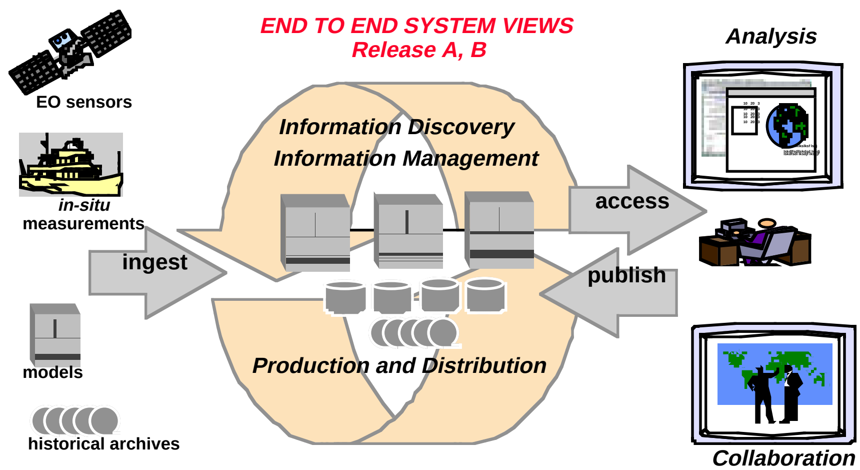


Figure 2-1. ECS Scenario Context

Figure 2-2 describes the SDPS in the context of its interfaces with other ECS segments and external systems. The key interfaces within ECS are with the Flight Operations Segment (FOS) to obtain Instrument Schedules and History Data, and with CSMS for status and coordination information as well as the full communications and system management infrastructure on which all SDPS applications are built. The primary external interface is with EDOS / ECOM to obtain EOS L0 data. The right hand side of the figure depicts other interfaces that SDPS must address.

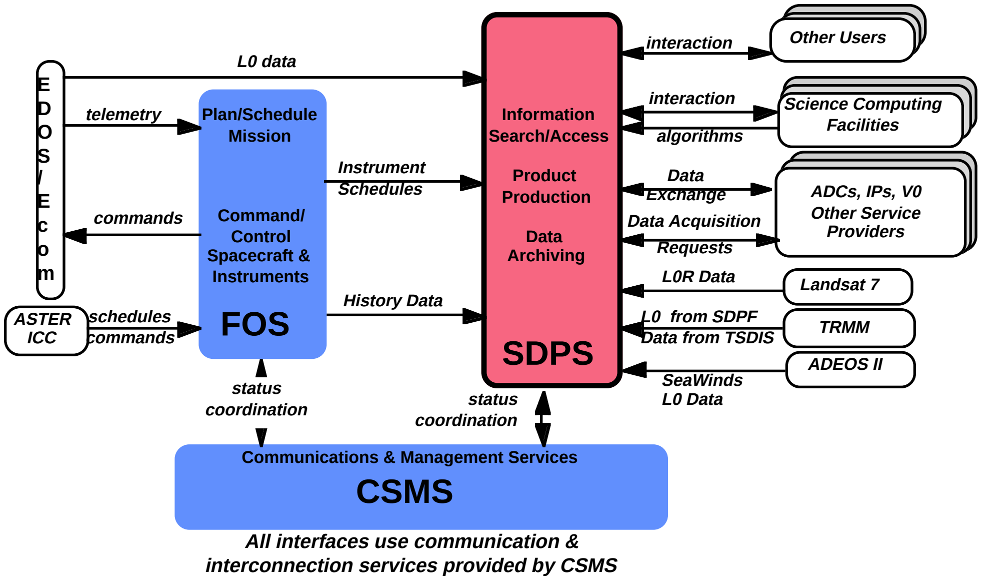


Figure 2-2. SDPS Segment Context

Figure 2-3 shows how all of the components in the conceptual architecture work together as an integrated system. Providers are depicted on the bottom. These components take raw data and provide for “ingest” into the system. Providers are geographically distributed, heterogeneous, and autonomously managed. They include not only the DAACs, but also the SCFs, ADCs and ODCs, and a set of secondary providers we call VAPs (value added providers). These providers will serve overlapping constituencies through a federated approach that is implemented in the distributed inter-site data management layer. The client layer includes core client environment support, custom applications, and user-specific COTS tools.

The federation of providers and customizable client layer support a diverse user community as depicted. The architecture provides capabilities in support of EOSDIS and GCDIS:

- Object- and service-oriented architecture
- Extensible client environment
- Evolutionary approach to Data Management
- Logical Distribution of components
- Transparent physical distribution and Service migration
- Site autonomy and heterogeneity
- Extensible provider network
- Domain specific representations
- Development of new types and services
- Incorporation of legacy systems
- Reusable components for GCDIS / UserDIS

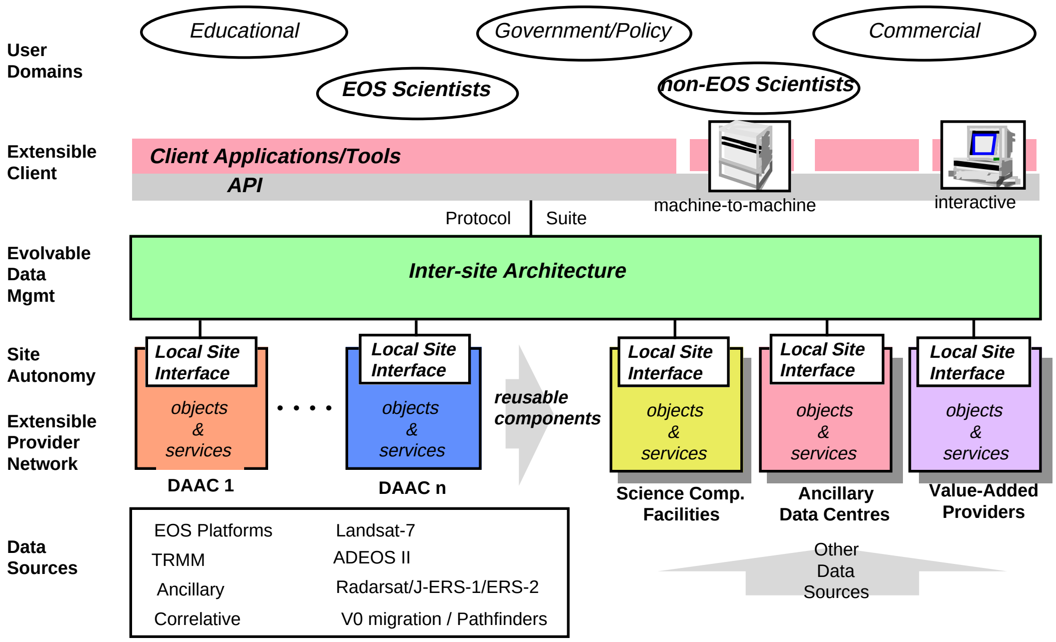


Figure 2-3. Conceptual Architecture

2.3 Role of Scenarios

The two key elements of a successful system design are the characterization of the static and dynamic elements of a system. In ECS the means by which these two aspects of a system are captured are System Scenarios, Object Models, and Component Interfaces.

System scenarios present the dynamic representation of the ECS as a system showing the system component interaction. Both an external user view and an internal system view are presented in each of the scenarios since the impression is that Earth Scientists are not interested in the internal system dynamics but rather what the system will do for them in conducting their research. The internal system view is important since the ECS developers at least need both a static object model and a dynamic scenario based model to gain an understanding of and to continue to refine and develop the internal representation and implementation of the ECS system.

The object model presents a static representation of the system, defining the science domain and system design vocabulary and the science and design concept relationships. The object model is defined in great detail in the System Design Specification (SDS) and the SDPS Segment/Element Design Specification.

The component interfaces also present a static representation of the system, defining the major system components, i.e. subsystems and Configuration Items (CIs) and the key interfaces among those components. The interfaces are linked to the object model via the services associated with each object and the key interface objects for complex interfaces.

2.4 Scenario Relationship to the User Characterization Model

The scenarios presented in this technical paper were derived from the formal user characterization efforts conducted in the ECS Science Office. The scenarios were further refined by direct collaboration with the contributing scientists. The focus of the scenarios for the SDPS PDR time frame was the functionality available at the Release A and B time frame. Similarly the user characterization effort is modeling the push and pull requirements for 4 epochs:

- January-June 1997 - Release A with a subset of the Version 0 data and activities supporting launch preparations;
- July-December 1997 - Release A, with TRMM;
- January-June 1998 - Release B, with TRMM and EOS-AM; and
- January-June 1999 - multi-mission support. The assumptions used in these models are described in a separate article (see "PDR Technical Baseline").

The User Characterization effort will estimate user demand for specific ECS services (service searching, metadata searching, browsing, subsetting/manipulation, data content searching, and data distribution). For the SDPS PDR, the emphasis was on EOS-funded investigators and non-EOS scientists, with coarse estimates for other user communities. The first two scenarios presented in this technical paper address the Earth Scientist pull on the system and the view of the Instrument Scientist and his team on the push side of the system.

Important inputs for user characterization, and subsequently the SDPS scenario development effort, come from the Version 0 User Services Working Group. Members from each V0 DAAC are compiling current usage statistics. These will form the starting point for projecting ECS usage at each of the 4 epochs defined above. This data will used to develop a pull - side generator for the performance model, representing the average user load, plus excursions for periods of peak demand. The data will also be used to refined the SDPS system scenarios.

As illustrated in Figure 2-2, the user characterization effort drove the model of the Scientists view of the system. In turn the scenarios link the user view with the internal system view.

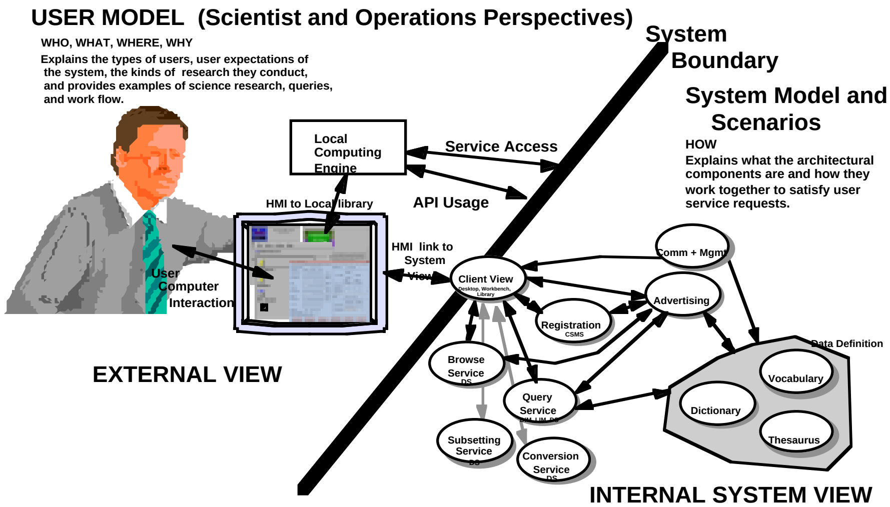


Figure 2-2. Scenario Relationship to the User Model

2.4 Scenario Relationship to the Data Model

The ECS system must eventually support a 15 Pbyte archive with hundreds of different data sets. The data model is a critical part of the design and is a driver for many of the design concepts. The data model is being developed in conjunction with representatives from throughout the EOSDIS community. The data pyramid based view of the data model is illustrated in Figure 2-3.

Four key features are included with the data modeling effort and help to simplify the design of the data management and data server subsystems.

Data Collections - a grouping of related data; e.g all products/information related to ocean dynamics (e.g product X and product Y would be in the same collection). These collections will be represented to users as logical collection within logical data servers (NOT to be confused with a physical data server described in the design). Collection groupings will change over time.

Data Pyramid - within a collection there are many types of related data with different characteristics and purposes. For the purposes of representation these are discussed in the context of a pyramid shown in Figure 2-3.

ESDT - the data type as 'viewed' by a user - e.g a MODIS SST product. Requests for data, actions on data will generally be expressed in terms of ESDTs.

CSDT - the physical representation of an ESDT in terms of component structures (e.g grid, image, text file, table, etc.). These are important - since it is at this level the code works, i.e. requested actions on ESDTs (e.g subset by parameter) are converted into functions on CSDTs. An ESDT may comprise one or more CSDTs.

Data Model is a critical part of the overall design

Key features . . . .

- **Data Collections**
- **Data Pyramid**
- **Earth Science Data Types [ESDT]**
- **Computer Science Data Types [CSDT]**

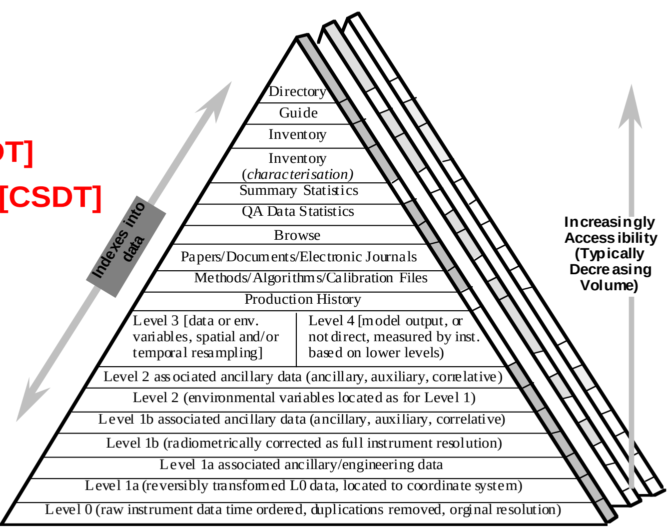


Figure 2-3. EOSDIS Data Model

2.6 SDPS External User View

The key components of the ECS that a user must interact with are contained within the Client subsystem and consist of the Desktop and Workbench. The Desktop provides the management and layout support for ECS specific graphical user interfaces running on a local research workstation. The Workbench supports users configuring their own local tools and provides management of the interfaces to distributed ECS tools and services. The service interfaces will be configured dynamically and will be managed locally via containers. As a result of invoking services, collections of results will be returned to the user's workstation for subsequent interaction. The returned results will be typed (i.e. are components of the system data and object models and are members of one of the data types defined in the object model) and can have services associated with them (either local or remote). This model for client server interaction applies not only to the GUI interface but also to machine-to-machine interaction.

Figure 2-4 portrays a mockup of the ECS Desktop and Workbench. The look and feel and content are intended to provide a conceptual understanding of the graphical user interface and how a user would view the system as a collection of data and services. The design of the Desktop and Workbench is an ongoing process and will be presented at several prototype workshops and design reviews for use in Release A and B over the next couple of years.

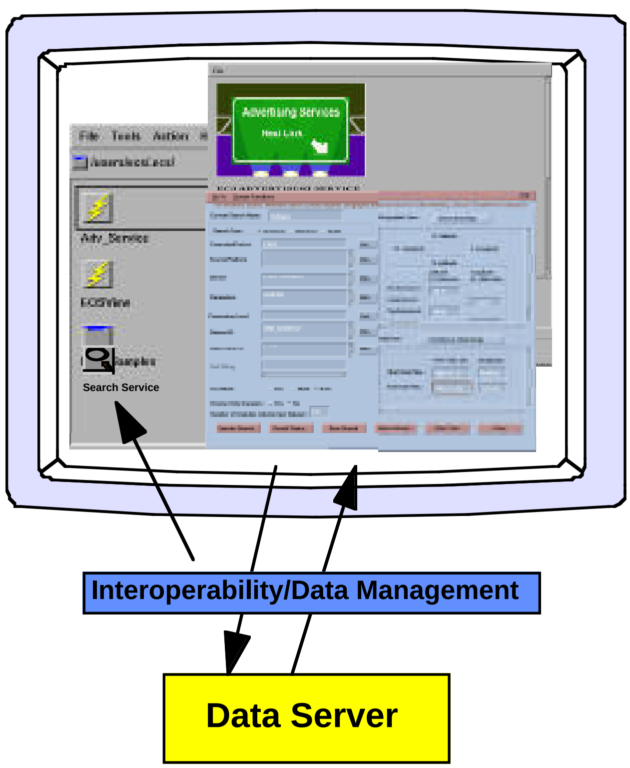


Figure 2-4. SDPS External View

2.6 SDPS Internal System Scenario Views

Figure 2-5 illustrates the seven subsystems of the SDPS and their interaction. The right side of each figure depicts (ingest, planning, and data processing) often described as the “PUSH” side of the system largely have to do with the “preparation” of data. This may be done at the time of ingest (eager) or request (lazy). Data is PUSHED into the Data Server. The left side of the figures (client, data management, interoperability/advertising) largely supports the information discovery components -- used for browsing, searching, and access / retrieval in a distributed environment. Data is PULLED from the Data Server. Note that in producing products the PUSH side must also act as a consumer of data and information from repository (Data Server), in this role the PUSH side components act as client and also PULL data from the Data Server. In the middle is the Data Server, which is the “information repository” or “information warehouse", provides the bridge between both pull and push sides of the system.

Figure 2-6 relates the three views presented in this document to each of the ECS SDPS subsystems. The PUSH side of the system is addressed primarily in the first scenario, Instrument Scientist view. The PULL side of the system is addressed primarily in the second scenario, Earth Scientist view. The third scenario addresses the middle context or the Data Server from an operations and administration view.

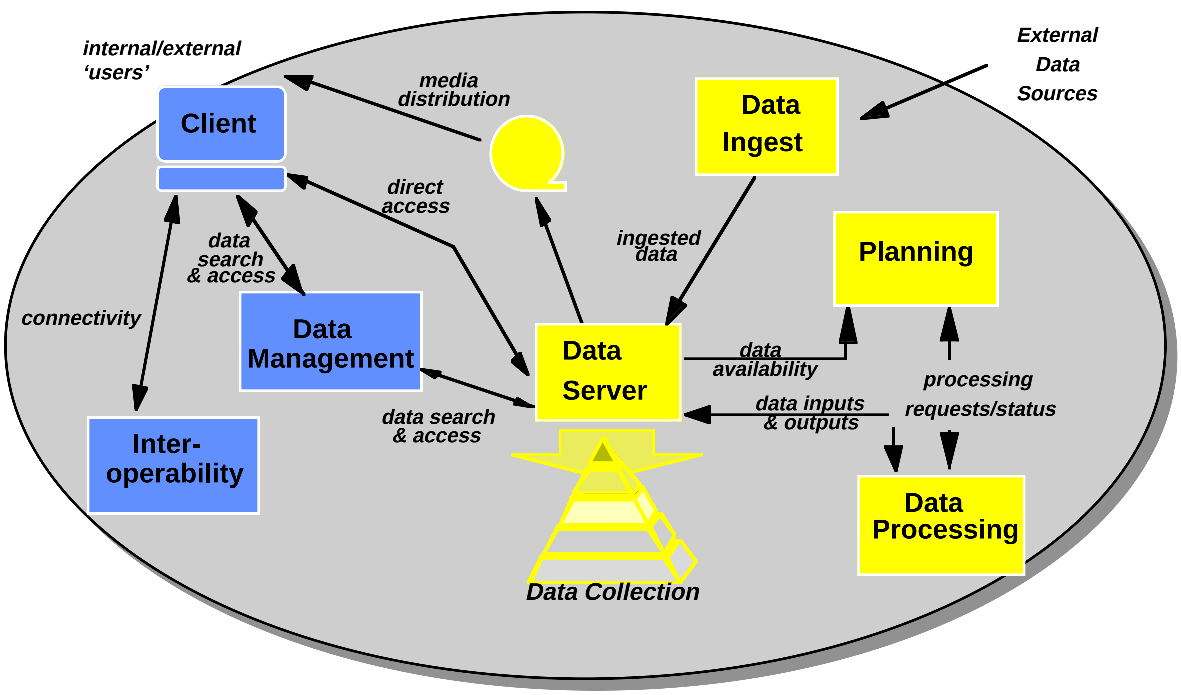


Figure 2-5. Subsystem Interaction Overview

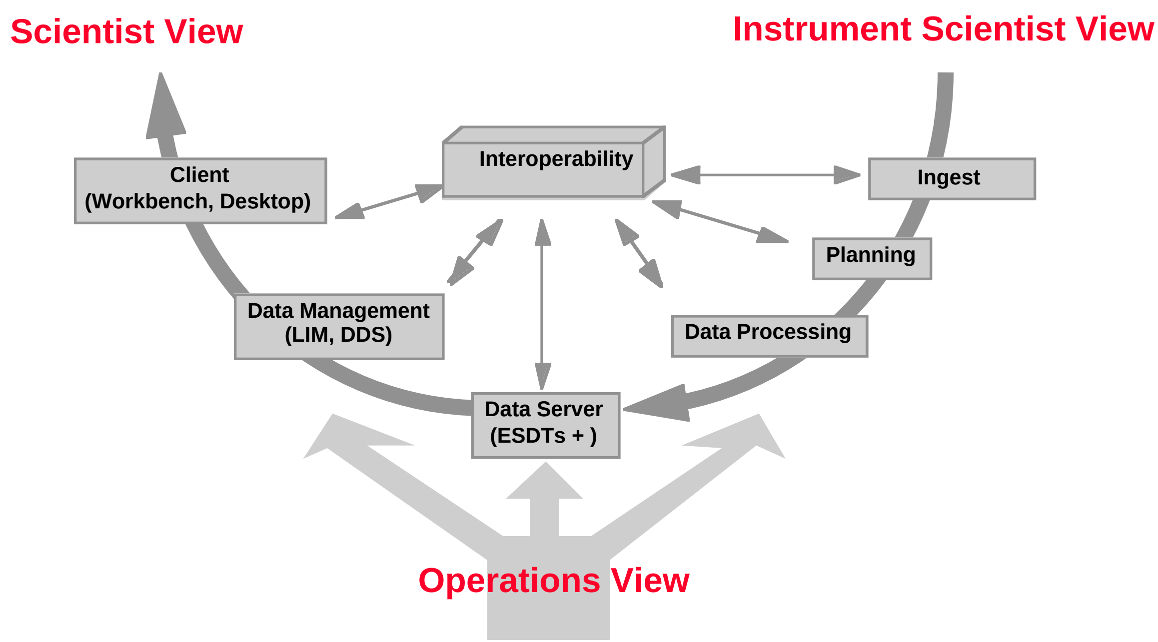


Figure 2-6. Internal System Scenario Views

2.7 SDPS End to End Scenarios

Three end to end scenarios are presented in the following sections providing a high level summary of the ECS system from three totally different points of view.

2.7.1 Instrument Scientist View

The first scenario is based on an instrument scientist's view of the system and covers the planning and data processing services. The scenario demonstrates the planning and processing steps for standard production and reprocessing in Release A and on-demand processing in Release B of CERES product algorithms.

2.7.2 Scientist View

The second scenario is based on a land processes researcher's view of the system during both the Release A and B time frames. DEM and AVHRR (Release A) /MODIS/Landsat-7 (Release B) data are combined to create surface models for the scenes of interest and validated with higher resolution ASTER data (Release B).

2.7.3 Administrative / Operations View

The third scenario is based on an administrator or operator's view of data server administration. The scenario demonstrates the planning, maintenance, and operations steps related to data server data type management.

2.8 Major Scenario Features of Releases A and B

The first scenario, CERES Product Planning and Processing, uses the Release A TRMM services of standard production, reprocessing, planning, and ingest and the Release B on-demand processing services. The second scenario, Access by a Land Processes User, uses the Release A V0 query, order, and browse interface, the ECS EOSView, workbench, desktop, advertising, data definition, and data server services and the Release B ECS query interface, local information manager, and data server subsetting services. The third scenario, Data Server Administration, uses the Release A Schema, Type Services, and Data Dictionary build services, the Advertising and data dictionary export services, and the test data load and access services. Release B data server administration services include the schema update and export services and the data server type service update services. Figure 2-7 lists each of the scenarios and associated ECS release functionality.

Scenario	Release A TRMM Services	Release B AM- 1 Services
A. CERES Product Planning	•Standard Production •Reprocessing •Planning •Ingest	•On-demand Processing
B. Access by Land Processes User	• V0 Query Interface • V0 Order • V0 Browse I/F • V1 EOSView (HDF) • V1 Workbench • V1 Desktop • V1 Advertising • V1 Data Definition • V1 Data Server	• V1 Query Interface • V1 LIM • V1 Subsetting
C. Data Server Administration	•Build Schema, DD •Build Type Services •Advertise, Export DD •Test data load / access	•Export Schema •Update Schema •Update Type Services

Figure 2-7. Major Features of the Releases

3. Instrument Scientist Scenario

3.1 CERES Planning and Processing Introduction

This scenario includes the planning process and the instrument team interaction, monitoring of the production status, and monitoring of the plan. The scenario covers the Release A and B scenarios for a CERES instrument scientist at LaRC. The scenario addresses the Earth's Radiation Budget science context, concentrating on the production processing and planning aspects of product generation. The scenario addresses the CERES instruments on the TRMM-1 mission in 1997 and the AM-1 mission in 1998, to gather information in support of long term monitoring of Earth's Radiation. The CERES instrument consists of two broadband, scanning radiometers: one cross-track and one rotating plane.

The CERES instrument extends the research and knowledge gained from the ERBE instrument. The CERES instrument measures the Earth's radiation budget and atmospheric radiation from the top of the atmosphere to the surface. The research will result in generating an accurate and self consistent cloud and radiation database. Cloud and radiation flux measurements are fundamental inputs to models of oceanic and atmospheric energetics, contributing to extended range weather forecasting.

The processing for the CERES products consists of 12 Subsystems:

- 1 Geolocate and calibrate Earth radiance,
  - 2 ERBE like inversion to instantaneous TOA and surface fluxes,
  - 3 ERBE like Averaging to Monthly TOA and surface fluxes
- 4 Determine Cloud Properties, TOA and Surface Fluxes
  - 9 Grid TOA and surface fluxes
  - 10 Monthly & Regional TOA and SRB Averages
- 5 Compute Surface and Atmospheric Radiative Fluxes
- 6 Grid single satellite radiative fluxes and clouds
- 7 Merge satellites, time, interpolate, compute fluxes.
- 8 Compute regional, zonal and global averages
- 12 Regrid Humidity and temperature Fields

The scenario illustrates the collaborative nature of the production processing and planning process between the CERES instrument team and the LaRC DAAC administration and operations staff. In the scenario the instrument team and operators provide inputs to the planning process and monitor the production status. The tasks illustrated in Figure 3-1 attempt to capture the parallel efforts between the DAAC and Instrument teams and specifically the coordination between the two teams. Three primary steps are illustrated in the scenario flow, first the analysis and generation of a plan with instrument team peer review, next the generation of an integrated plan for the LaRC DAAC, and finally the review and monitoring of the resulting plan and schedule. Eight specific steps are identified as part of the scenario and are described in the next section.

3.2 Scenario System Flow

The eight steps, listed in Figure 3-2, consist of (1) the DAAC staff enters a standard production request and the production scheduler is notified. The (2) Data Availability Schedule (DAS) is available and is viewable via the ECS provided planning services. The (3) production scheduler initiates creation of candidate plans for review and (4) the production scheduler (and, in parallel, (6) the instrument team) views the plans. Once the plan is accepted and data is available the plan is implemented and a schedule for execution is developed (in coordination (7) with the instrument and DAAC operations staff teams) and executed. During execution (5) the production scheduler (and (8) instrument teams) view(s) the PGE profile and monitors the progress of production at the LaRC DAAC.

3.3 Scenario Design Overview

The scenario is supported by four key ECS subsystems: Ingest, Data Server, Data Processing and Planning, as illustrated in Figure 3-3. The key features of the design in support of the scenario are listed in Figure 3-4. A common user interface supports the instrument team in providing inputs to the planning process. The ECS design will provide support in resolving differences in naming conventions (if any exist) between the project standards and the instrument team standards. The design will also support multiple ingest push and pull models through the use of Ingest Clients specialized for each external interface. The configuration management associated with the planning process is automated, thereby minimizing the amount of manual intervention. A consistent view of the plans and schedules, schedule for product release, and any plan/schedule updates is maintained and made available to authorized users. The toolkit interface to the data server services is consistent with the design of the data server and the algorithm integration and test framework. Parallel planning for quality assurance and algorithm integration and test activities is supported by the ECS design.

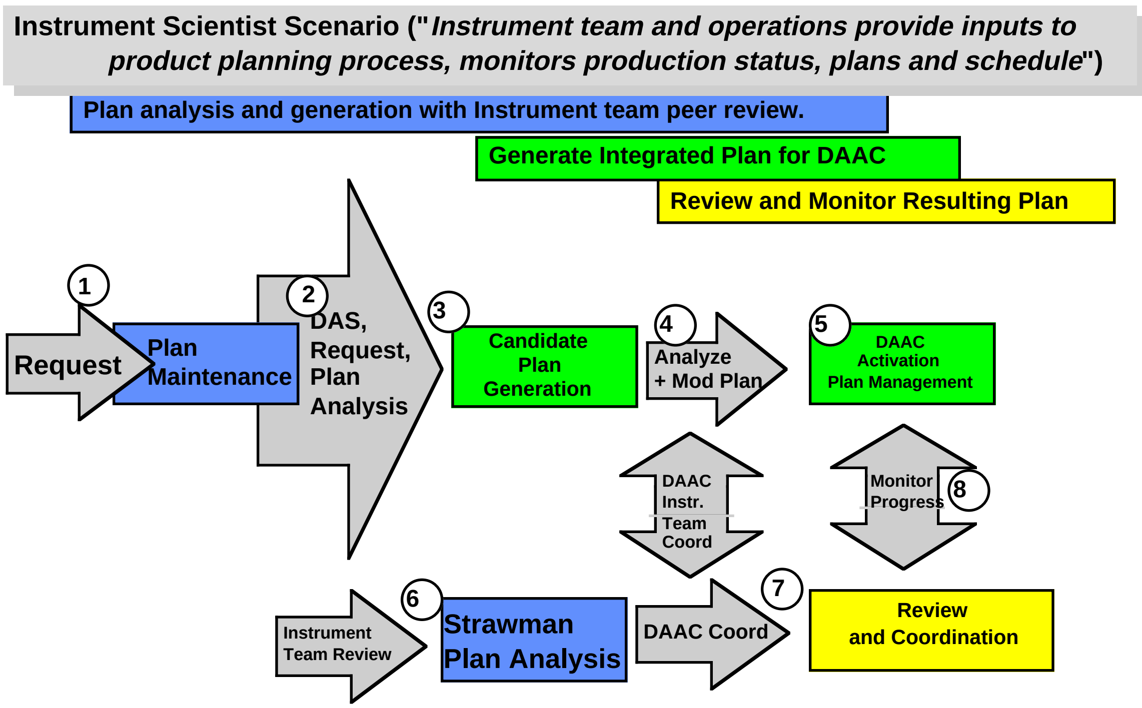


Figure 3-1. Instrument Scientist Scenario Overview

1. User Enters Standard Production Request and Production Scheduler Notified
2. DAS for SDPF Available and Viewable
3. Production Scheduler Initiates Creation of Candidate Plan
4. Production Scheduler Views Plan
5. Production Scheduler Views PGE Profile
6. Instrument Team Reviews Strawman Plan
7. Instrument Teams and DAAC Operation Staff Coordinate Plan Activation
8. DAAC Operations Staff and Instrument Team Monitor Progress

Figure 3-2. Scenario System Flow

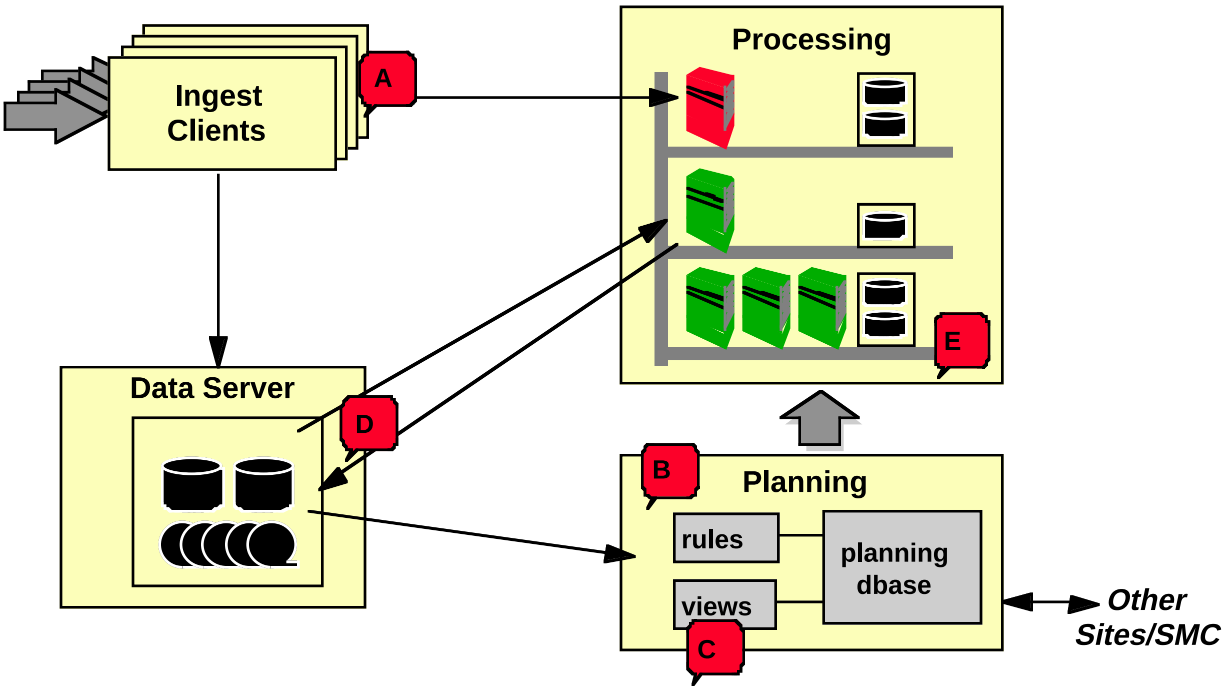


Figure 3 -3. Scenario Science Overview

A Instrument Team Provides Inputs to Planning Process
- Naming Convention Differences
- Ingest Push and Pull Models

B Configuration Management of Planning is Automated

C Consistent Views of Plans Maintained
- Schedule for Product Release
- Plans and Schedules
- Plan and Schedule Updates

D Toolkit Interface to Data Server Services is Consistent with Design

E Parallel Planning (QA) and Processing (AI&T) Activities

Figure 3-4. Key Design Features

4. Earth Scientist Scenario

4.1 Access by a Land Process User Introduction

The scenario described in this section is derived from the User Model Scenario #7 titled " The development of a method to integrate datasets of varying resolution". The three primary steps in the scenario are the development of a technique to integrate data from sensors of varying spatial, temporal, and spectral resolutions, the generation of a surface model, and the verification of the model with independent inputs.

The scenario is based on the research approach developed by Dan Baldwin that applies state of the art information technology to AVHRR reflected channel data to create a "super resolved" model of surface albedo and derived NDVI composites. The results can be used to determine surface type heterogeneity within an AVHRR pixel, which is important for computation of surface to atmosphere fluxes.

Repeat 1km AVHRR data are non-coincident-- centers of repeat FOVs wander over a specified grid box, so that each repeat measurement contains different information from features within the grid. Bayesian maximum posterior estimates can be used to deconvolve differences from multiple repeat AVHRR data sets and reconstruct the surface data at a resolution higher than individual AVHRR measurements.

The first step in the method is to use a high resolution (180m) DEM in conjunction with solar and satellite viewing geometry to create a constant albedo model of the satellite measured illumination. The next step is to, register FOVs from multiple AVHRR data sets to the model grid and integrate model values over AVHRR point spread function to give model estimate of reflectivity. Bayesian theory is applied to determine corrections to the constant albedo values that best estimate AVHRR measured reflectivities. The result is reconstructed albedo on 180m grid. Finally the method verifies the results with Landsat TM data. The primary assumption of the method is that the surface characteristics are static over accumulation period of data.

Two stages were identified as part of the scenario definition: the development stage, using Release A functionality and data; and the EOS era production stage using Release B functionality and datasets. The development stage begins with the selection of the development region-- Requirements are : 1) availability of high resolution DEM, 2) relatively static surface features, 3) moderate topography, 4) infrequent cloudiness. Next the researcher must determine the accumulation period -- based on availability of cloud free Landsat TM scenes for verification and the availability of cloud free AVHRR data, close to nadir. The region selected for the scenario is 1 x 1 degree region in Death Valley, time period 04/01/92 - 06/01/92.

The data requirements for the development stage include, 90 meter DEM (later degraded to 180m);2 cloud free TM scenes over the region, preferably in early April and late May, 1992; 20 AVHRR data sets, Level 1 or 2, reflectivity or albedo, bands 1 and 2, minimal cloud contamination, within 20 degrees scan angle of nadir, early afternoon pass; and any other data describing surface characteristics of region: albedo, surface types, vegetation, etc.

The EOS era production stage uses the same research strategy but substitutes MODIS data instead of AVHRR. The composite product is verified with the high resolution ASTER data. Figure 3-1 represents pictorially the development stage flow using AVHRR browse products to determine the coverage of the area of interest, the AVHRR products, subsetted at the researchers site, and subsetted DEM data to generate a higher resolution composite product.

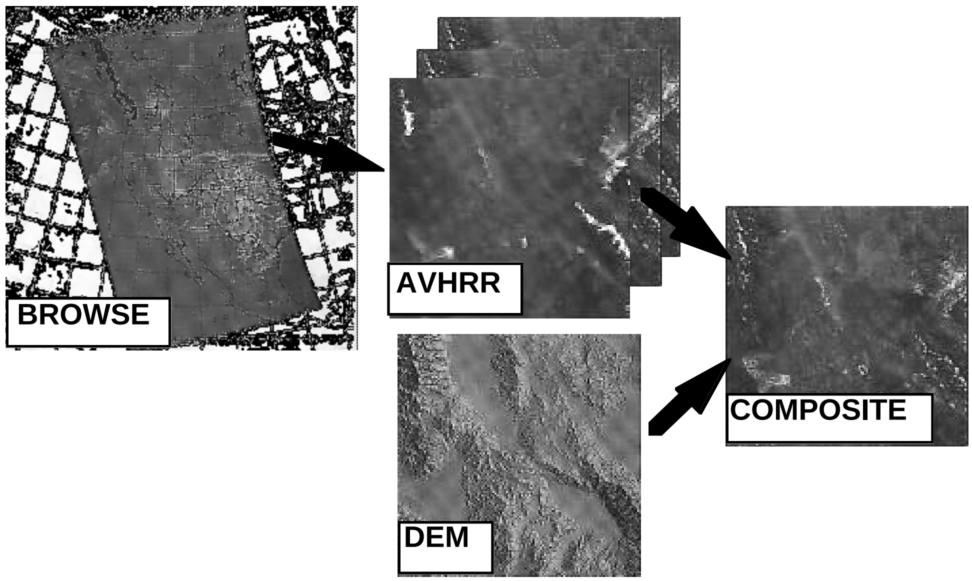


Figure 4-1. Earth Science Research Overview -- Release A

The Release B scenario, depicted in Figure 4-2, follows the same flow described in the introduction above substituting the EOS era products and subsetting functionality. The flow is depicted as fifteen steps, each of which is described in the next section.

4.2 Scenario System Flow

The scenario flow is itemized in Figure 4-3 and includes the step by step description of the science scenario in user terms. The flow begins with a (1) Search Service access to be used for Landsat-7 data. The results are returned and a selection is made by the researcher. Next (2) the user views the query results and selects the appropriate granules. The researcher then (3) inspects the images and (4) requests the retrieval of the selected images (note in Release A the entire granule is returned (5) and in Release B the subsetted image is returned (5) to the researcher's site). The researcher then (6) searches advertisements for DEM providers and (7) retrieves the DEM fragments of interest . The researcher again, using the advertising services (8) searches for MODIS Level 1B search service providers and links to the search service. Once linked the researcher (9) searches for the MODIS data in the time period, geographic area, and parameter constrained subset required. The MODIS data is (10) received at the researchers local workstation and the researcher (11) analyzes the Landsat, MODIS, and DEM products to create a surface model at his own workstation. Once the product and process has been tested and integrated with the MODIS provider site, the user (12) defines a private ‘user method’ at the local workstation (Release A + B). Subsequently the researcher (13) issues a standing request to obtain data for 10 day intervals (Release A+B). The researcher (14) establishes a standing request for cloud-free ASTER scenes and then (15) receives the requested data and stores in local archive.

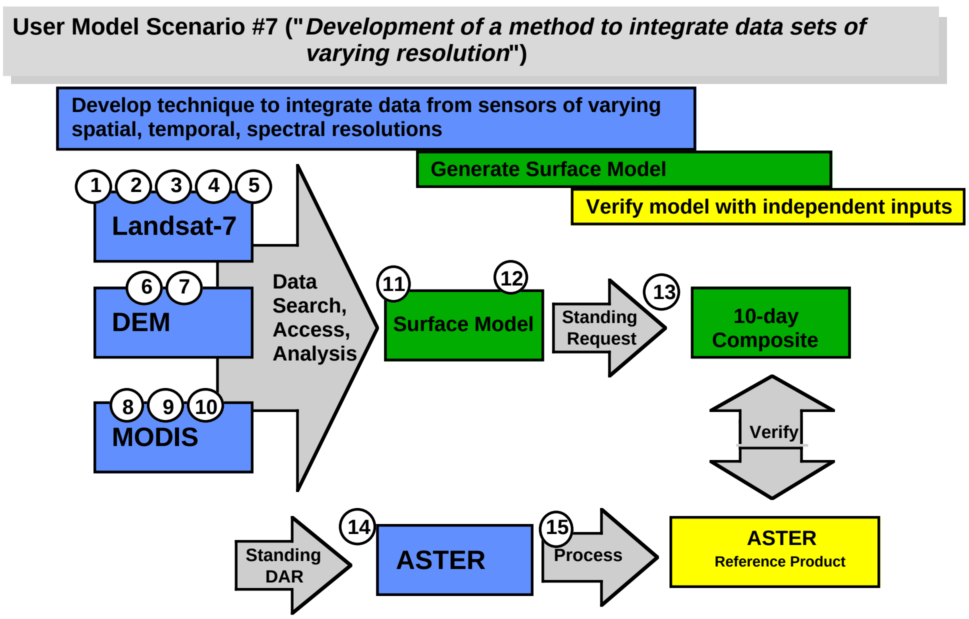


Figure 4-2. Earth Scientist Scenario Overview

1	Search Service Used for Landsat-7, Results Returned & Selection Made
2	User Views Query Results, Selects Granules
3	User Inspects Image
4	User Request Full granule of each selected image
5	Data Received at Local Workstation
6	Search Advertisements for DEM Providers
7	Retrieve DEM Fragments and Deliver to User
8	User using Advertising to Access MODIS Level 1B Service Provider
9	Search for MODIS Data
10	MODIS Data Received at Local Workstation
11	User Analyzes Landsat, MODIS and DEM to Create Surface Model
12	User Defines private ‘user method’ at local workstation (Release A + B)
13	User issues a standing request to obtain data for 10 day intervals (Rel A+B)
14	User establishes request for cloud-free ASTER scenes
15	User receives requested data and stores in local archive

Figure 4-3. Scenario Science Flow

4.3 Scenario Design Overview

The key subsystems involved in this scenario are depicted in Figure 4-4 and include the Client, Data Management, Interoperability, and Data Server. The key features relevant to the Release A and B designs are listed in Figure 4-5.

Key design features for Release A include the use of the desktop and workbench in the client subsystem to integrate the ECS and V0 services and the advertising service supporting both ECS and V0 services. Release A user interfaces for search, browse and retrieval are based on the V0 interface. Release B user interfaces are based on ECS developed advertising, search, browse, and retrieval.. During both the Release A and B time frames the Data Server provides access to type services (both Earth Science Data Types and Computer Science Data Types) for search and access interfaces. In addition, at the Release B time frame, the Data Server will provide subsetting for the requested products.

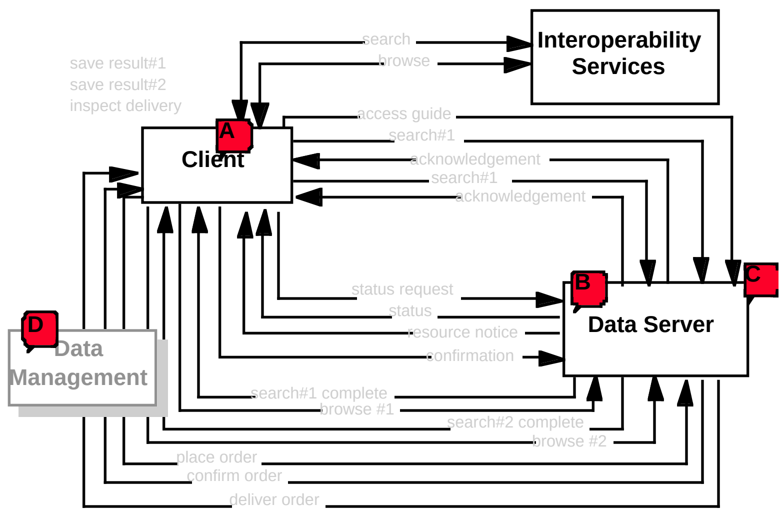


Figure 4-4. Design Overview

A **V1 Desktop and Workbench Integrates ECS and V0 Services**
- **Search, Advertising, and Data Definition Services**
- **Data Server Type Services for Search & Access at Release A**
- **Data Server Type Services for Subsetting at Release B**

B **V0 Search Service used to access V0 servers and ECS Data Servers**

C **WWW access to Document Data Server**

D **Integrated Site View at Release B**

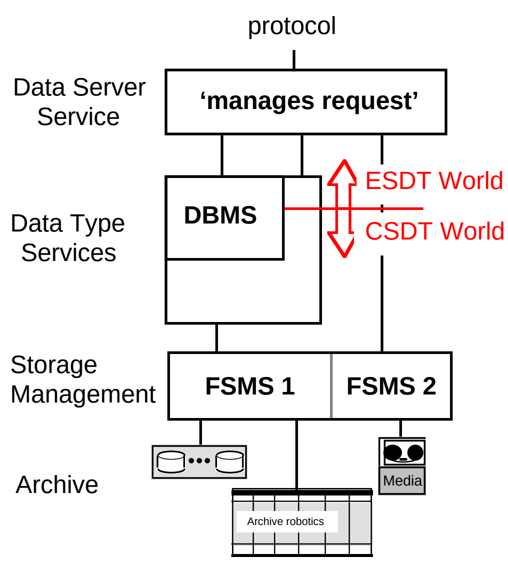
Figure 4-5. Key Design Features

5. Administrative / Operations Scenario

5.1 Data Server Administration Introduction

This scenario includes the functionality necessary, during the Release A and B time frames, to support data server administration for the CERES product specific model (documented in the CERES data modeling documents and the data server design documents.) The scenario describes the steps to develop a data server model, load test data and subsequently verify operations. A more detailed description of these activities is included in the Data Server Subsystem specification documents.

The data server design, illustrated in Figure 5-1, provides a separation between the logical design components and the physical implementation through the use of Earth Science Data Types (ESDTs) and Computer Science Data Types (CSDTs). Through this design each data server can implement services in any of a number of technologies and can support evolvability and extensibility through the definition of new types and services. The implementation strategies of the data server include the request protocols associated with each of the data types such as documents, images, tabular data, and complex three-dimensional data sets. The role of the data management systems as implemented via Relational, Extended Object Relational, and pure Object Oriented database management systems is encapsulated within the ESDT layer. Archive and archive access technologies (such as distributed file systems and hierarchical storage management systems) are encapsulated within the CSDT layer.



Manages data for both the PUSH and PULL aspects of EOSDIS.

Key Points

- **Receives requests in terms of ESDTs which are converted to functions on CSDTs**
- **DBMS selection influences design solution**
- **must support multiple archive technologies**
- **highly desirable that it supports multiple FSMS**
- **highly desirable that it supports sub-file access**

Figure 5-1. Data Server Design

The data server exports its services and data collection structure via two complementary mechanisms: advertising, data dictionary export and schema export. As illustrated in Figure 5-2 the advertising mechanism links the data server services with the rest of the system via the Interoperability subsystem. Advertisements contain a description of the service being offered (such as search through the CERES products), the location of the service, a reference to the service that can be retained by clients, and searchable keywords characterizing the service. The data dictionary exports contain information defining the terms, valid values, relationships, and synonyms managed by each data server. The schema defines the structure of the data maintained by the data server and is exported to the Local Information Manager (LIM) for integration with the local site view of all of the data holdings. The LIM exports the site view, again as a schema based description, and this information is unioned by the Distributed Information Manager (DIM) into a federated view of the entire system. The DIM can then support system wide, cross-site distributed queries. The federated schema maintained by the DIM allows for global optimization of query process, query decomposition and allocation to relevant sites, and result integration of partial responses from each of several sites.

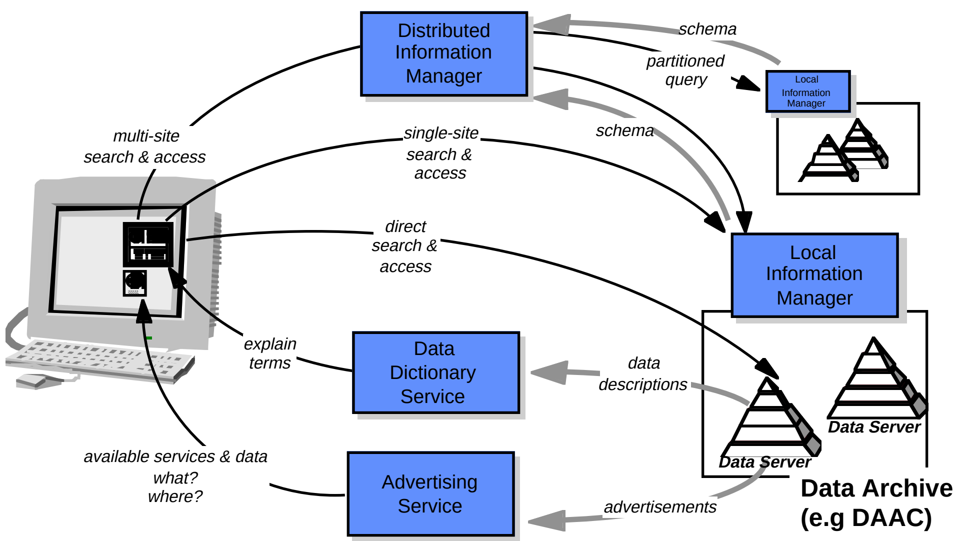


Figure 5-2. Data Server Interaction with Interoperability and Data Management

Conceptually, from the scientist's perspective, the Data Server appears to be a collection of data objects and the services available on them. But, from the implementation or physical perspective, the data server does not exist as a single physical component. Rather, it is a collection of software, hardware, operations and data that support a collection of data objects and access to them.

The Data Server stores and provides access to many different data types. Just as the data server has a logical and a physical view, so does the data that it stores. From the data server perspective, the logical view is called the Earth Science Data Type (ESDT) and the physical view is called the Computer Science Data Type(CSDT). Figure 5-3 illustrates the collection of ESDTs for the CERES data model and the associated primary CSDT. The CER03 ESDT is further refined into the primary Rectangular Grid data type, and secondary Access Table and Science Data Table data types. Each CSDT is implemented using the Science Data Processing Toolkit utilities developed to support the HDF storage and distribution format.

The ESDT represents what is known to the user community and includes the various levels of product data, browse data, guide data, etc. The CSDT represents what is known to the physical computers that store and manipulate the data and includes arrays, lists, binary trees, etc. This bioptic view has always existed whenever computer programs processed real world data. The distinction is made to attempt to provide a common reference between the users of the data and the implementers that must ensure that the data is stored and accessed properly.

ESDTs and CSDTs are implemented by combining off the shelf DBMS technology with software developed to support unique ECS data requirements that are not supported by the off the shelf software directly. For example, the DBMS will support simple data types such as integer, floating point, string, arrays, images, time, and various kinds of spatial objects. However, when building database support for ECS products, software may need to be developed to support coordinate transformations which are not supported by the off-the-shelf software.

Information which resides in a DBMS is defined formally by a database schema. The DBMS uses the schema to interpret queries and access requests. The data objects, i.e., the ESDT, are defined in the schema and are not visible in the software design of the Data Server itself. The schema would implement the requirements of the ECS Core Meta Data Model (with product unique extensions) developed by the Data Server design team and the Data Modeling team. The Core Meta Data Model would not be reflected in the software design, only in the database schema.

Maintenance and operation of the Data Server involves the management of the collection of ESDT's, CSDT's, implementation libraries, schemata associated with the data types, advertisements, data dictionary entries, exceptions and associated events, subscriptions, archive file systems, and the supporting hardware.

The steps associated with this scenario are illustrated in Figure 5-4. This scenario describes how to set up a data server and how to build and integrate the schema for a new data server into the data management subsystem. Major activities include: implementation of the data server administration and operations services to develop a data server model; loading of test data and verification of operational status; generation of the data model components; and complete testing of the data server functionality.

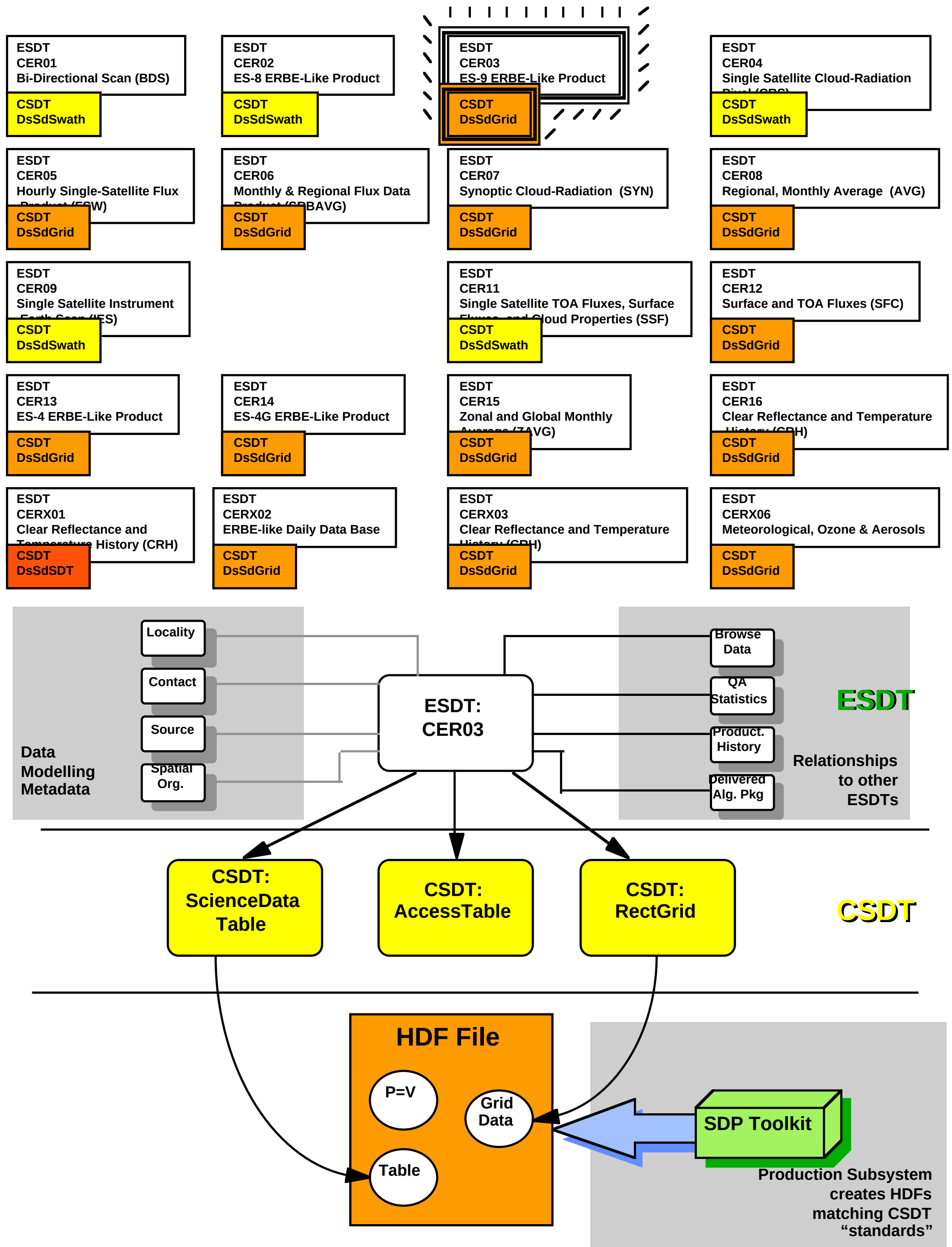


Figure 5-3. Data Server ESDT and CSDT Design for CERES Release A

5.2 Scenario System Flow

Six major steps, illustrated in Figure 5-5, are associated with the operations scenario for the data server and include (1) the administrator receives a request for a new data server and gathers the characterization and test data for the new data server. Next (2) the administrator checks the validity of the data type definitions, the consistency of the schema, and determines any site unique extensions that must be maintained by the data server. Then (3) the administrator analyzes the schema and data type views, creates an internal and export schema definition, advertises the services and data collections in a test mode, and sets up the links among the DBMS's, file systems, and archive systems. Next (4) the administrator checks for completeness and consistency of the data server logical and physical data models. After successful completeness and consistency checking, the administrator (5) issues a request for the operator to bulk load test data to verify the content and operational state of the data server. Finally (6) the operator and administrator runs a full operational test on the integrated data server components. After successful completion of the test the data server is re-initialized and prepared for standard production quality inputs.

5.3 Scenario Design Overview

Four sub-systems are involved in the administration of the data server, the Client, Interoperability, Data Management and the Data Server sub-systems, as illustrated in Figure 5-6. Three key design features, listed in Figure 5-7, are relevant to the administration and operation of the data server. First the ECS design provides online support for the administration and operation of the data server via a configurable user interface that provides common access to all administration, operations, and production services. Second the design supports the integration of the data server services with the data management and advertising services. This support includes the export of schema to the Local Information Manager and the subsequent integration and test within the LIM; the support for export and test of data dictionary term definitions, valid values, and synonyms; and the generation and test of advertisements of data server services and data collections. Third the design provides support for data server operational testing including tools to support loading of test data, and access to completeness and consistency test utilities.

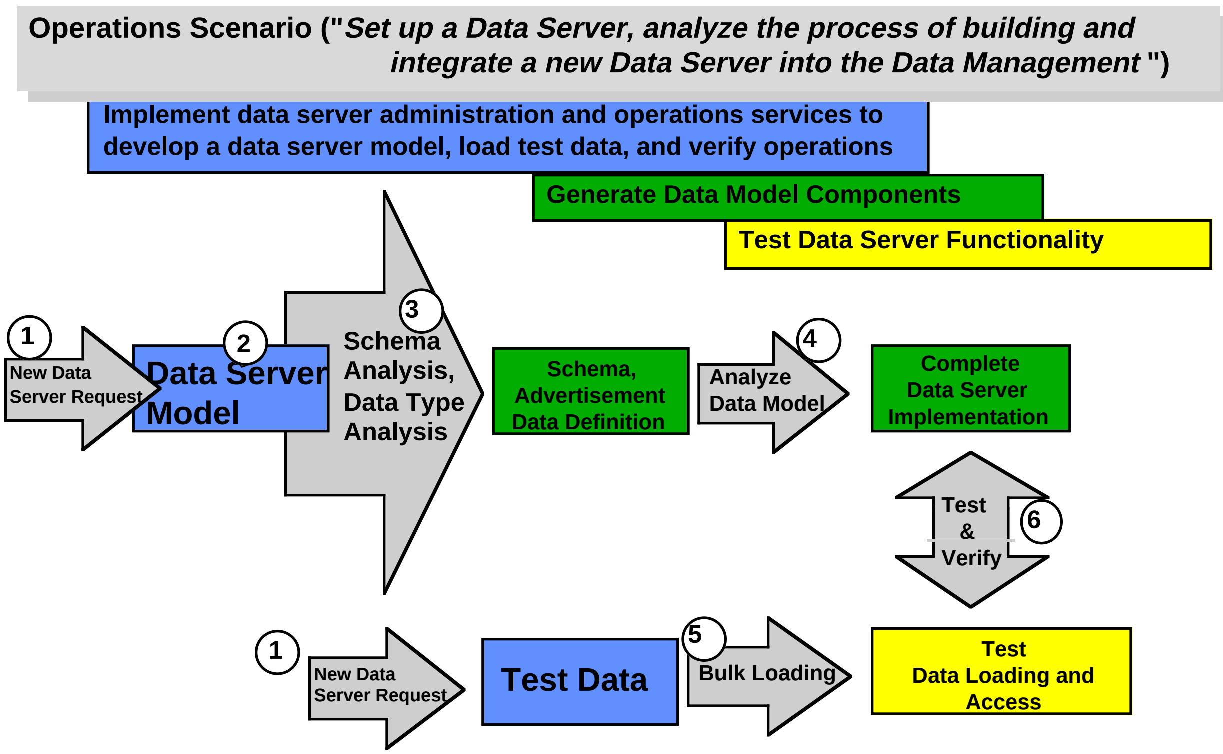


Figure 5-4. Administration and Operations Scenario Overview

1. **Administrator Receives New Data Server Request and Gathers Data Server Characterization and Test Data**
2. **Administrator Checks Validity of Data Type Definitions, Schema Consistency, Site Unique Extensions**
3. **Administrator Analyzes Schema and Data Type Views and Creates Internal and Export Schemas,Advertisements, and DBMS, Files System and Archive Linkages**
4. **Administrator Checks Complete Data Server Logical and Physical Data Models for Consistency and Completeness**
5. **Operator Bulk Loads Test Data and Verifies Content of Data Server**
6. **Operator and Administrator Integrates and Tests Data Server Components for Full Operational Test**

Figure 5-5. Scenario Flow

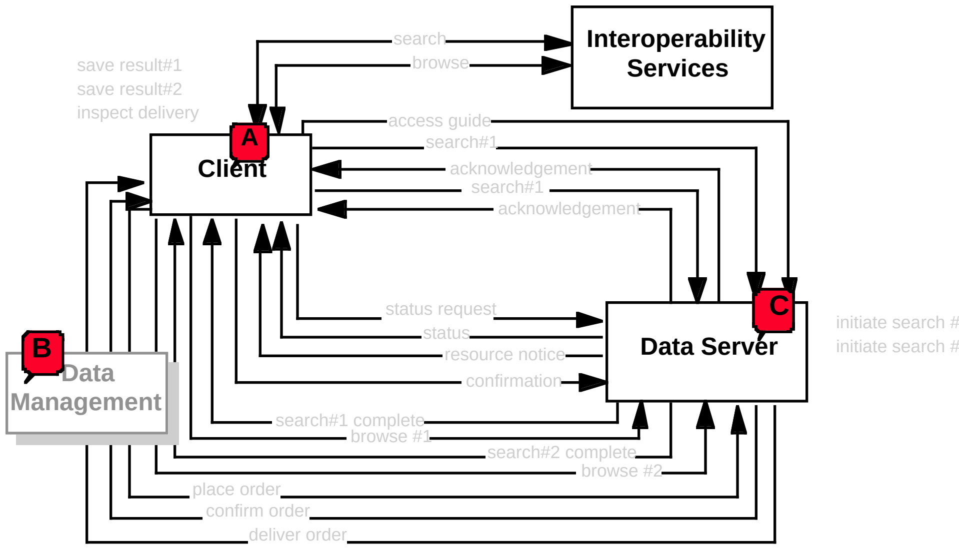


Figure 5-6. Design Overview

A Online Support for Administration and Operations Services
- **Configurable User Interface**
- **Common Access to all Services**

B Online Support for Data Management & Advertising Integration
- **Schema Export, Integration and Test within LIM**
- **Data Dictionary Term Definition Export and Test**
- **Advertisement Generation and Test**

C Online Support for Data Server Operational Testing
- **Bulk Loading of Test Data**
- **Data Access Test Utilities**

Figure 5-7. Key Design Features

6. Suggested Reading

Asrar, G. and Dokken, D., 1993: 1993 EOS Reference Handbook. NASA Headquarters. URL: http://spso.gsfc.nasa.gov/eos_reference/TOC.html

Asrar, G. and Dozier, J, 1994: Science Strategy for the Earth Observing System. NASA Headquarters.

Corbell, M., 1994: MR9407V3 User Environment Definition for the ECS Project.
URL: http://edhs1.gsfc.nasa.gov/

Dopplick, T., 1995: 160-TP-003-001, A Science User's Guide to the EOSDIS COre System (ECS) Development Process. URL: http://edhs1.gsfc.nasa.gov/

ECS, 1994: ECS Prototyping and Studies Plan. URL: http://edhs1.gsfc.nasa.gov/

ECS, 1994: Summary of the ECS System Design Specification. URL: http://edhs1.gsfc.nasa.gov/

ECS, 1995: 305-CD-002-002, Science Data Processing Segment (SDPS) Segment/Element Design Specifications for the ECS Project. URL: http://edhs1.gsfc.nasa.gov/

ECS, 1995: 301-CD-002-003, System Implementation Plan for the ECS Project (1/95).
URL: http://edhs1.gsfc.nasa.gov/

Elkington, M., 1994: 19400131 Defining the Architectural Development of EOSDIS.
URL: http://edhs1.gsfc.nasa.gov/

Elkington, M., Meyer, R. and McConaughy, G., 1994: Defining the Architecture of EOSDIS to Facilitate Extension to a Wider Data Information System. Proceedings of the ISPRS'94, Ottawa.

Endal, A., 1994: FB9402V2 ECS Science Requirements Summary - Version 2.
URL: http://edhs1.gsfc.nasa.gov/

EOS Instrument Principal Investigators, 1994: Algorithm Theoretical Basis Documents. URL: http://spso.gsfc.nasa.gov/atbd/pg1.html

James, M. 1995: EOSDIS Data Set Reference Handbook. In preparation, to be published in Spring 1995 by NASA/GSFC.

Moxon, B. 1993: 19300611 Science-based System Architecture Drivers for the ECS Project.
URL: http://edhs1.gsfc.nasa.gov./

Moxon, B., 1994: FB9401V2 EOSDIS Core System Science Information Architecture.
URL: http://edhs1.gsfc.nasa.gov/

NASA, 1994: EOSDIS VO IMS Home Page.
URL: http://harp.gsfc.nasa.gov:1729/eosdis_documents/eosdis_home.html/

National Research Council, 1994: Panel to Review EOSDIS Plans, Final Report. National Academy Press.

Schwaller, M. 1993: Science Data Plan for EOSDIS Covering EOSDIS Version 0 and Beyond. NASA/GSFC. Anonymous ftp host: eos.nasa.gov. Change directory to /EosDis/Daacs/Docs/ScienceDataPlan.

Tyahla, L. 1994: 19400313 ECS User Characterization Methodology and Results.
URL: http://edhs1.gsfc.nasa.gov/

Tyahla, L. 1994: 19400548 User Scenario Functional Analysis. URL:
http://edhs1.gsfc.nasa.gov/

Unninayar, S. and Bergman, K., 1993: Modeling of the Earth System in the Mission to Planet Earth Era. NASA Headquarters.

West, S., 1994: 19400549 ECS Scientist User Survey (ESUS). URL: http://edhs1.gsfc.nasa.gov/

Wingo, T. 1994: 19400312 User Characterization and Requirements Analysis.
URL: http://edhs1.gsfc.nasa.gov/

This page intentionally left blank.

Acronyms

ADC	Ancillary Data Center
AVHRR	Advanced Very High-Resolution Radiometer
CERES	Clouds and Earth's Radiant Energy System
CI	Configuration Item
COTS	Commercial off the Shelf
CSDT	Computer Science Data Type
CSMS	Communications and System Management Segment
DAAC	Data Active Archive Center
DBMS	Data Base Management System
DEM	Digital Elevation Model
DIM	Distributed Information Manger
ECS	EOSDIS Core System
EDOS	EOS Data and Operations System
EOSDIS	Earth Observing System Data and Information System
ERBE	Earth's Radiation Budget Experiment
ESDT	Earth Science Data Type
FOV	Field of View
FSMS	File Storage Management System
GCDIS	Global Change Data and Information System
LIM	Local Information Manager
MODIS	Moderate-Resolution Imaging Spectroradiometer
NDVI	Normalized Difference Vegetation Index
ODC	Other Data Center
PB	Petabytes(10^15 bytes)
PDR	Preliminary Design Review
QA	Quality Assurance
SCF	Scientific Computing Facility
SDPS	Science and Data Processing Segment

SDR	System Design Review
SDS	System Design Specification
SST	Sea Surface Temperature
TB	Terabytes (10^12 bytes)
TM	Thematic Mapper
TOA	Top of the Atmosphere
TRMM	Tropical Rain forest Monitoring Mission
URL	Uniform Resource Locator
UserDIS	User Data and Information System
VAP	Value Added Provider
WWW	World Wide Web